Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

ALDH1L1 SNPs and their allelic combinations are associated with glycine-and purine-related metabolites in a family-based cross-sectional study

Sergey A Krupenko
University of North Carolina at Chapel Hill
Sai Sravani Vennam
University of North Carolina at Chapel Hill
Lauren Roppolo Brazell
University of North Carolina at Charlotte
Faustina Therase Jeyaraj
University of North Carolina at Chapel Hill
Kendra L. Nelson
University of North Carolina at Chapel Hill
Brea C. Nance
Standard Process
Chinmayee Panda
Standard Process
Baba B. Mass
University of North Carolina at Chapel Hill
Venkata Saroja Voruganti
saroja@unc.edu

University of North Carolina at Chapel Hill

Research Article

Keywords: Family-based study, Non-synonymous variants, Metabolomics

Posted Date: November 25th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11653120/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.

Abstract

Background: The gene encoding folate enzyme, cytosolic aldehyde dehydrogenase 1 family member L1 (*ALDH1L1*) has been shown to regulate glycine and purine metabolism. Our aim in this study was to determine whether exonic non-synonymous single nucleotide polymorphisms (SNPs) in *ALDH1L1,* and their genotypic and allelic combinations are associated with glycine and purine metabolites in a family-based study, the NRI Children Health Study (n = 409, 116 families). We genotyped six *ALDH1L1* exonic SNPs: rs1127717, rs2276724, rs2886059, rs3796191, rs4646750 and rs9282691(MAF: 4 – 22%). Glycine and purine metabolites were assessed as part of an untargeted metabolomics platform. We tested the influence of each SNP, genotypic and allelic combination on metabolites via a regression-based analysis, considering their kinships. Results; The most significant SNP association was observed for sarcosine with rs4646750 [β (SE) = -0.27 (0.06), p = 9.7 E-6), myo-inositol with rs1127717 [β (SE) = 0.28 (0.07), p = 6 E-5) and histidine with rs3796191 [β (SE) = 0.14 (0.03), p = 3.2 E-5]. For genotypic combinations, the most significant association was found between adenosine/inosine ratio and rs1127717*TT_rs2276724*CT_rs2886059*AC_ rs3796191*AG_

rs4646750*TT_rs9282691*TT[β (SE) = 0=-191 (24.5), p = 5.9 E-14], and histidine and rs1127717*TT_

rs2276724*TT_rs2886059*CC_rs3796191*AG_rs4646750*TT_rs9282691*TT [β (SE) = 0.22 (0.04), p = 2.1E-7]. Further analysis with allelic combinations showed several unique combinations linked to glycine and purine metabolites. The most significant association being that of adenosine/inosine ratio with rs2276724*C_

rs3796191*G[β (SE) = -638.4 (95), p = 7.4 E-11] and histidine and rs2886059*A_rs3796191*A [β (SE) = -0.32 (0.07), p = 9.9 E-6]. Conclusions: These results reflect the additive effect of genotypic and allelic combinations of *ALDH1L1* exonic non-synonymous SNPs on glycine-and purine-related metabolites with implications for related metabolic diseases.

INTRODUCTION

The ALDH1L1 enzyme catalyzes the $NADP^+$-dependent conversion of 10-formyltetrahyrdofolate (10-fTHF) to tetrahydrofolate (THF) and carbon dioxide.[1,2] This reaction controls the availability of one-carbon groups for folate-dependent metabolic processes including *de novo* purine biosynthesis.[2,3] It also replenishes the cellular THF pool, which is required for the conversion of serine to glycine, and the degradation of glycine and histidine. ALDH1L1 has been viewed as a candidate tumor suppressor based on its downregulated expression in malignant tumors.[1–4] Studies have linked loss of *ALDH1L1* to enhanced tumor progression in lung adenocarcinoma,[5] hepatocellular carcinoma (HCC),[6] pilocytic astrocytoma[7] and breast cancer.[8] Higher *ALDHL1* expression was also significantly associated with reduced risk of recurrence in patients with HBV-related HCC.[9] In agreement with a tumor suppressor role, *ALDH1L1* is silenced in most cancer cell line through methylation of its promoter region while ALDH1L1 expression in those cells evokes strong antiproliferative effects.[10,11] ALDH1L1 has also been implicated in the etiology of other pathological conditions including aortic insufficiency,[12] neural tube defects,[13] Alzheimer's disease, [14,15] and non-alcoholic fatty liver disease.[16]

ALDH1L1 expression and its effect on metabolism can depend on specific genetic variants defined by common single nucleotide polymorphisms (SNPs).[2] Genetic studies have shown SNPs in *ALDH1L1* are significantly associated with serum concentrations of folate, glycine, serine and glycine-to-serine ratio.[13,17–19] In a previous study from our group, exonic SNPs in *ALDH1L1* were associated with these metabolites, and also with BMI, waist circumference, HDL cholesterol, IL6 and total homocysteine along with the purine metabolism end product, urate.[17] The association of *ALDH1L1* SNPs with homocysteine was also observed in participants of Framingham study.[20] This study further found that *ALDH1L1* SNP rs2364368 is associated with incident of ischemic stroke.[20] In a trans-omics study of 618 adult participants, variants in *ALDH1L1* were linked to diabetic kidney disease.[21] In a large cohort comprising of four studies with participants of European ancestry found rs1107366 in *ALDH1L1*, but not other SNPs, to be associated with insulin sensitivity.[19]

These studies show the importance of *ALDH1L1* gene and its variants in cancers and other diseases. However, association of single SNP by itself may not be informative with respect to health status and may not explain the total effect of the group of

SNPs on the phenotype.[22–25] Therefore, we aimed to study the role of *ALDH1L1* exonic SNPs, their genotype combinations and the allelic combinations in the variation in glycine and purine metabolites. To accomplish the aim, we used a family-based cohort with 409 participants, comprising of children and their parents from 116 families.

METHODS

Study population and sample collection

NRI Children Health study (NCHS) is a family-based cohort study aimed at identifying genetic factors affecting obesity in children. This study includes children, and their parents from Cabarrus and Rowan counties of North Carolina. The inclusion criteria were broad: children between 5 and 19 years of age if at least one parent willing to participate in the study, and without any ascertainment for sex, education, ethnicity, socio-economic status or disease. There was only one exclusion criterion: children or adults could not participate without the other. This analysis comprised of 2000 relative pairs from 116 families, with majority of them being first and second cousins followed by parent-offspring and sibling relationships. All adults gave written informed consent. Parental permission was obtained for all children, and assent was obtained from children 15–17 years old. All protocols (#16-1826) were approved by the Institutional Review Board at the University of North Carolina at Chapel Hill in accordance with the Common Rule (under 45 CFR 46).

Anthropometric and blood pressure measurements

Height was measured using a portable stadiometer. Body composition and weight were obtained using a body composition analyzer (TBF-300A, Tanita Corporation, Arlington Heights, IL), with subjects shoeless and wearing light clothing. Waist circumference (WC) was measured between superior iliac crest and costal margin. All anthropometrics were measured in duplicates, and the means of the measurement were used for analysis. Blood pressure was measured using the Omron digital blood pressure monitor (HEM907XL, Omron Healthcare Inc, Lake Forest, IL) at least three times following a 5-min rest. The means of the last two measurements were used for analysis. Body mass index (BMI) was calculated using the formula "(Weight in kilograms/[Height in meters]^2)."

Metabolomic profiling (Glycine and purine metabolites)

Glycine, purine and related metabolites were measured as part of the untargeted metabolomic profile. Nontargeted metabolomic profiling was performed using plasma samples from NCHS and replicate samples from MTRX3, a large, extensively characterized pool of human plasma maintained by Metabolon (Durham, NC, USA).[17,26–29] Briefly, 3 independent platforms were used: ultra-HPLC-tandem MS optimized for basic species, ultra-HPLC-tandem MS optimized for acidic species, and GC-MS. The original raw ion intensities for all metabolites were normalized to account for the instrument inter-day tuning differences. Largely, each compound was corrected in run-day blocks by registering the medians to equal 1.00 and normalizing each data point proportionately (block correction). In certain conditions, to account for differences in metabolite levels due to variation in the amount of material present in each sample, biochemical data were normalized to addition factors. Values that were missing or below the limits of detection were imputed with the compound minimum values. All metabolite values are scaled, and imputed values of their intensity are presented as means ± SDs. The final data comprised 1305 metabolites, after excluding metabolites that were missing for more than 40% of participants. We selected 27 metabolites for this analysis based on their links to glycine and purine metabolism.

DNA extraction and Genotyping

DNA was extracted from buffy coat via an automated nucleic acid extraction platform (Anaprep 12, Biochain, Institute Inc. Newark, CA, USA) using the Anaprep Blood DNA extraction kit (Biochain Institute Inc. Newark, CA, USA). Concentration and purity of genomic DNA was measured using NanoDrop Spectrophotometer. Six *ALDH1L1* exonic SNPs: rs1127717, rs2276724, rs2886059, rs3796191, rs4646750 and rs9282691, were genotyped. Genotyping was performed via Taqman® predesigned SNP

genotyping assay (Applied Biosystems, Foster City, CA, USA) on Quant Studio 12k Flex Real-time PCR system (Thermo Fisher Scientific, USA).[30,31]

Statistical analysis

Baseline characteristics of the study population are expressed as percentages for categorical variables, and as the mean and standard deviation. Additionally, baseline metabolic measurements (not yet log transformed) are expressed for the subset of metabolites chosen for this study that impact the glycine and purine metabolism, but were log-transformed to achieve normal distributions for all subsequent analyses. The six SNPs were tested for Hardy-Weinberg equilibrium, as well as for the presence of linkage disequilibrium.

To consider the genetic relatedness of the individuals, pedigrees and kinship matrices were generated from the participants in the study. A variance components decomposition method was used to estimate heritability.[32] This method is implemented in the software program SOLAR and has been described in detail elsewhere.[32] All hypotheses were evaluated using standard likelihood ratio test procedures and the associated likelihood ratio test statistic (LRT). The LRT is computed as minus twice the difference in ln-likelihoods estimated under the null and alternative hypotheses and is in standard cases distributed as a $\chi 2$ with degrees of freedom (d.f.) given by the difference in the number of estimated or unconstrained parameters. However, as is often the case for variance components models, the non-standard case of testing a null hypothesis lying at a boundary of its acceptable parameter range yields a LRT distributed as a 50:50 mixture of a point-mass at 0 and a $\chi 2$ with 1 d.f. Prior to conducting genetic analyses, distributional properties of all traits were evaluated. All values beyond four standard deviations of the mean were removed and the remaining traits were transformed by inverse normalization prior to analysis, to meet assumptions of normality. Then, the effect of individual SNPs on metabolic measurements were tested by a regression-based association test, taking into account the age, sex, and age*sex of each individual. For the SNP associations, statistical significance was set at $p < 8$ E-3 after adjusting for multiple tests.

Frequencies among the 6 SNPs of interest were translated into distinct genotypic combinations represented using their nucleotides. Genotype combinations with frequencies > 2% were used in analyses that follow. Standard linear models were fit for each metabolite, using the distinct combinations as response variables, and similarly, using age, sex, age*sex, an individual's status reflecting whether they were a child/parent, and each metabolite as a predictor covariate. Additionally, multinomial regressions were performed to determine the probability of differences. The analyses were performed by using base R, the HaploView package in R, and the SOLAR software on UNC's high-powered research computing cluster. The statistical significance for analysis of genetic combinations was set at $p < 7$ E-4 after adjusting for multiple tests (SNP combinations plus phenotypes).

We conducted allelic combination analyses to investigate the effect of multiple SNPs on metabolic traits. Analyses were performed in R (RStudio version 2025.05.1_513). For each trait, we systematically generated 2-way through 6-way combinations of alleles and modeled their combined effects using linear regression. Each allelic combination was treated as a categorical predictor in the model. (presence of a combination vs all others). From these models, we derived overall p-values using ANOVA to assess whether the allele combination as a whole was significantly associated with the trait. In addition, we obtained combination-specific regression coefficients (β), standard errors, and p-values, which provide insight into the direction and magnitude of effect for each allelic pattern. This analysis is adapted from the analytical methods previously described by Jung et al.[33] To support interpretation, we also calculated the number of individuals carrying each combination, allowing us to distinguish between effects driven by common combinations versus those arising from rare variant combinations. This analysis allowed us to evaluate both pairwise and higher-order allele interactions across traits. The statistical significance for allelic combinations was set at $p < 7.8$E -5 after adjusting for multiple tests (allelic combinations plus phenotypes).

RESULTS

The study comprised of 116 distinct families and included 39% male and 61% female participants, with a mean (SD) age of 23.3 (15.5) years. A total of 30 phenotypes were selected for the analysis. They include 27 metabolites from the untargeted metabolomic panel based on their links to glycine or purine metabolism and 3 ratios of metabolites as potential surrogate

markers of the enzyme activities (Inosine/hypoxanthine – purine nucleoside phosphorylase, adenosine/inosine – adenosine deaminase and serine/glycine - serine hydroxymethyltransferase) . Their scaled intensities of select metabolites and ratio values are presented in **Table 1**. Next, we examined whether the serum metabolites concentrations were heritable. The heritability of a phenotype indicates the effect of additive genetic factors on the variation in phenotype. Significant additive genetics (heritabilities) were observed for all 27 selected metabolites ($p < 0.05$) (**Table 1**).

Six exonic non-synonymous SNPs of *ALDH1L1* were genotyped in this study. **Table 2a** presents details about the six SNPs including their allele frequencies. None of the studied SNPs deviated significantly from the Hardy-Weinberg equilibrium. Their minor allele frequencies ranged between 4% and 22%. Genetic association analysis of these metabolites with the *ALDH1L*1 exonic SNPs showed associations with 8 metabolites that maintained significance after multiple testing adjustments ($p < 0.008$). The strongest associations were observed for rs4646750 with sarcosine ($p = 9.7 \times 10^{-6}$), rs3796191 with histidine ($p = 3.2$ E-5) and for rs1127717 with a glycine conjugate, myo-inositol ($p = 6.1$ E-5)(**Table 2b**).

A total of 40 genotypic combinations were identified with the six SNPs. The combinations for which frequencies were more than 2% were included in the analysis. The top *ALDH1L1* genotypes based on the combination of the six SNPs are shown in **Table 3**. About one third participants had the combination of rs1127717*TT_rs2276724*TT_rs2886059*CC_ rs3796191*AA_rs4646750*TT _rs9282691*TT. Others were present in 15% or less of the participants. The combination of genotypes were tested for association with the relevant metabolites. The most significant associations were observed for adenosine/inosine ratio with rs1127717*TT_rs2276724*CT_rs2886059*AC_rs3796191*AG_ rs4646750*TT_rs9282691*TT ($p = 5.9$E14), rs1127717*TT_rs9282691*TT ($p = 5.9$E14), rs1127717*TT_ rs2276724*TT _rs2886059*CC_

rs3796191*AA_rs4646750*TT_rs9282691*GT($p=1.7$E-06), and for histidine with rs1127717*TT_ rs2276724*TT_rs2886059*CC+_rs3796191*AG_ rs4646750*TT_ rs9282691*TT ($p = 2.1$E-07). Those who had the first combinations had significantly lower values of adenosine/inosine compared with others [β = −191.005 (24.529) and −123.824 (25.497), respectively], whereas those who had the second combination had significantly higher concentrations of histidine compared with others [β = 0.22 (0.04)]. **(Table 4).** Other metabolites that were significantly linked to *ALDH1L1* SNP combinations were dimethylglycine and inosine. Individuals with rs1127717*TT_ rs2276724*TT_rs2886059*AC_

rs3796191*AA_rs4646750*TT_rs9282691*GT had higher levels of dimethylglycine as compared to those who didn't have this combination. On the other hand, individuals with rs1127717*TT_ rs2276724*TT_rs2886059*AC_ rs3796191*AA_rs4646750*CT_ rs9282691*TT had higher concentrations of inosine than those who didn't have this combination (**Table 4**).

A total of 607 unique allelic combinations were generated for the six *ALDH1L1* SNPs. For each trait,

we systematically generated 2-way through 6-way combinations of alleles and modeled their combined effects using linear regression. Of these, 41 combinations were significantly associated with the metabolites at a p value of less than 8.2 E-5 after adjustment for multiple tests. The most significant association was of adenosine/inosine ratio with rs2276724*C_rs3796191*G [b (SE) = -638.4(95.4), $p = 7.4 \times 10^{-11}$] and rs1127717*T_rs2276724*C_rs3796181*G [b (SE) = -645.2 (96.5), $p = 7.7 \times 10^{-11}$]. Histidine, cinmoylglycine, pantothenate, adenosine and inosine also showed significant associations with various other allelic combinations ($p < 8.2$ E-5). (**Table 5**).

DISCUSSION

Most genetic studies focus on the identification of SNPs within genes to determine their relation to alterations of the circulating metabotype. Depending on the extent of the metabolic changes, their effect on human health and predisposition to metabolic diseases can vary significantly. One of the genes in folate metabolism, *ALDH1L1*, has 6 common non-synonymous exonic SNPs which lead to the expression of the enzyme with altered amino acid sequence.[2] ALDH1L1 enzyme is involved in the regulation of the intracellular folate pool, in the metabolism of glycine and in *de novo* purine biosynthesis.[34–36] Therefore, in the case of altered functionality of the enzyme due to an amino acid substitution, the effect on cellular metabolism could be expected. In fact, our previous study of a cohort of Hispanic children has linked two *ALDH1L1* exonic SNPs to altered plasma concentrations of serine and glycine and the serine to glycine ratio.[17] These SNPs were also associated with changes in certain parameters, commonly

used as characteristics of metabolic health. Of note, some of these exonic ALDH1L1 SNPs are genetically linked together yielding 10 common ALDH1L1 haplotypes.[2]

In the present study we aimed to analyze the effect of combinations of genotypes and alleles rather than single SNPs on plasma metabotype in a local family cohort. Our study includes a family-based cohort of 116 families with family sizes ranging between 2 to 6. Family-based studies provide the best opportunity to identify variants with large effect sizes.[37] They not only share a genetic background but also share many household-level environmental exposures that minimize noise in detecting genetic variance. These studies also limit genetic heterogeneity as alleles can be tested for segregation within families.[37–39] They also provide advantage of better-quality control of the genotype data as they can be checked for Mendelian inconsistencies.[40,41]

We analyzed 27 metabolites in the two pathways, glycine and purine, and 3 ratios of metabolites as surrogate markers of enzyme activities. All the selected metabolites were significantly heritable. Heritability of a phenotype, metabolite in this case, refers to the additive genetic effects influencing the metabolites.[42] Previously, we and others have shown that these metabolites are heritable.[17,43] A study in more than 8000 individuals from the Canadian Longitudinal Study on Aging (CLSA) have found significant heritabilities for glycine, serine, urate and nucleosides such as inosine and xanthine.[44] In a study of siblings, Underwood et al. (2025)[45] found histidine to be significantly heritable (69%) in 2 cohorts comprising of 502 sibships. Our association study showed significant relation of histidine, myo-inositol, and N-formylmethionine with *ALDH1L1* SNPs. These findings are in agreement with the role ALDH1L1 has in these pathways.[35] Previous studies have implicated ALDH1L1 in the progression of several types of malignant tumors (reviewed in Krupenko and Krupenko 2020).[3] It has also been shown that the hypermethylation of the *ALDH1L1* promoter region suppresses expression of the gene in many cancers.[10] While *ALDH1L1* methylation has been established as the mechanism downregulating this gene in cancers, the effect of non-synonymous SNPs on the gene functionality is less obvious. One of such SNPs, rs2276724, has been linked to reduced expression of ALDH1L1 mRNA in hepatocellular carcinoma.[9] However, the effect of *ALDH1L1* SNPs is not limited to cancer. Our previous study have shown that these SNPs are linked to a range of metabolic and disease characteristics/markers such as anthropometrics, serum lipids, inflammatory markers and homocysteine.[17] Similarly, other studies have shown role of such SNPs in ischemic stroke,[20] coronary artery disease,[46] and diabetic kidney disease.[21]

Since it is known that a single SNP may not explain the complete genetic variance, we generated combinations of genotypes of each exonic SNP with other such SNPs. Our best association for genotypic combination was for adenosine/inosine ratio with the combination of rs1127717*TT, rs2276724*CT, rs2886059*AC, rs3796191*AG, rs4646750*TT and rs9282691*TT: the carriers of this combination tended to have lower values of this ration. We computed this ratio as a surrogate marker for adenosine deaminase (ADA) which catalyzes the reaction of adenosine to inosine conversion. Adenosine deaminase is an intracellular and cell-surface enzyme widely expressed.[47,48] Although there is no direct molecular link between ALDH1L1 and ADA, ALDH1L1 may influence its substrate adenosine via regulation of *de novo* purine biosynthesis and one-carbon flux. [2,3] Histidine, an essential amino acid, also showed significant association with rs1127717*TT_rs2276724*TT_rs2886059*CC_ rs3796191*AG_

rs4646750*TT_rs9282691*TT. It plays a major role in neurotransmission, renal function, gastric secretion and immune function.[49] It also has anti-inflammatory and antioxidant properties.[50] ALDH1L1 plays an important role in regulation of folate metabolism including folate-dependent histidine breakdown. Specifically, histidine metabolism is dependent on ALDH1L1 for regeneration of the THF pool necessary for histidine degradation.[49,51] Thus, alterations of ALDH1L1 functionality can affect the conversion of histidine to glutamate, the pathway requiring THF in its final step, which utilizes formiminoglutamic acid (FIGLU) to yield glutamic acid and to donate one-carbon group to the reduced folate pool.[49,52] In mouse models, pathway analysis based on untargeted metabolomics showed increased levels of the intermediate of histidine metabolism FIGLU in *ALDH1L1* KO mice.[35] The other key glycine metabolite, dimethylglycine, was also significantly associated with the genotype combination. Individuals with the combination showed greater concentration of dimethylglycine (DMG) as compared to those who didn't. This metabolite is produced in the pathways of homocysteine remethylation by betaine, which is linked to choline metabolism. DMG itself is catabolized to glycine in two sequential reactions, each donating a methyl group to THF. Thus, through production of glycine and one-carbon groups, DMG participates in the synthesis of neurotransmitters and hormones. Previously, genome-wide association studies (GWAS) have reported significant association of dimethylglycine with SNPs in dimethylglycine dehydrogenase (*DMGDH*),

the enzyme transferring the methyl group from DMG to THF thus linking it to one-carbon metabolism.[53,54] Consistent with these findings, studies of *aldh1l1* KO mice have demonstrated that the loss of ALDH1L1 leads to DMG accumulation, which is in agreement with insufficient THF in these mice.[35,55,56] Thus, the association of certain *ALDH1L1* genotype combinations with higher levels of dimethylglycine is indicative of insufficient production of THF, which in turn suggest decreased ALDH1L1 functionality.

The other two metabolites that were significantly associated with ALDH1L1 genotype combinations were serine and inosine. Serine, a glycine precursor, has been associated with *ALDH1L1* variants in previous studies of humans[2,17] and mice.[2] However, there are no previous studies linking inosine to *ALDH1L1* SNPs. Since inosine is a purine and ALDH1L1 has a role in *de novo* purine biosynthesis, the association assumes significance. This also shows that the differences in patterns or combinations of SNPs in *ALDH1L1* may have a role in variation in glycine and purine metabolites.

To extend our knowledge of the role of *ALDH1L1* genetic variants, we analyzed the patterns or combination of alleles of the six *ALDH1L1* SNPs. We identified 607 unique allelic combinations, with 41 of them showing significant associations with the metabolites studied. Most significant associations of allelic-patterns were with purine-related metabolites, adenosine, inosine and their ratio. ALDH1L1 functionality is also linked to the regulation of purine biosynthesis due to competition for the same substrate, 10-fTHF. It has been proposed that the enzyme controls the number of one-carbon groups available for the purine pathway.[3,55] Inosine, also produced as a microbial metabolite, is a breakdown product of adenosine catalyzed by ADA.[57,58] ALDH1L1 plays an important in donating carbons at two important steps in formation of the purine ring in generating inosine monophosphate (IMP) which is the precursor of AMP, GMP and inosine.[55,57] Although, no previous studies have the examined the relation between ALDH1L1 SNPs and adenosine and/or inosine, the role of ALDH1L1 in IMP generation and the indirect effect on ADA makes the link plausible.

Our study has few limitations. The sample size is relatively small and may suffer from false positive results. Future studies are needed to replicate and validate our findings. And, secondly, we used cross-sectional data in the analysis, and the results may not reflect causal relationship. However, our study also has some strengths. First, the family-based structure of the study provides greater power to detect meaningful associations among unique allelic combinations and glycine metabolites, and secondly, the selection of biological meaningful, targeted and known functional variants of *ALDH1L1* reduce the burden of multiple-testing and increase the likelihood of detecting true associations.[59,60]

In conclusion, genotypic and allelic combinations of *ALDH1L1* SNPs were associated with glycine and purine metabolites. This study captures the cumulative effect of the exonic SNPs and their alleles accounting for greater percent of variance and adds more evidence linking ALDH1L1 to purines and glycine metabolic pathways. This also study provides in-depth analysis of genotypes and alleles of the exonic SNPs going beyond the individual SNP effects and provides a closer insight into the relationship of ALDH1L1 with two metabolic pathways.

Declarations

FUNDING SOURCES

This study was partially funded by a NIH/NIDDK grant DK126666 (SAK &VSV) and NIH/NIDDK-P30DK056350, and partially by funds from the University of North Carolina at Chapel Hill Nutrition Research Institute (VSV).

HUMAN ETHICS AND CONSENT TO PARTICIPATE DECLARATION

All adults gave written informed consent. Parental permission was obtained for all children, and assent was obtained from children 15–17 years old. All protocols (#16-1826) were approved by the Institutional Review Board at the University of North Carolina at Chapel Hill in accordance to the Common Rule (under 45 CFR 46).

Author Contribution

SAK and VSV conceived and conducted the research. VSV supervised study design, analysis, interpretation and manuscript preparation. SSV and LRB contributed to the analysis and interpretation of the data. FTJ, KLN, BCN, CP and BBM contributed to the study population recruitment, implementation of study procedures, collection of data and genotyping. VSV and SAK obtained funds for the study.

Acknowledgement

We thank the participants (children and their parents) for their generous participation without which this study would not be possible.

Data Availability

The dataset supporting the conclusions of this article is available from the corresponding author (VSV) on reasonable non-commercial request.

References

1. Krupenko SA, Krupenko NI. ALDH1L1 and ALDH1L2 Folate Regulatory Enzymes in Cancer. In: Vasiliou V, Zakhari S, Mishra L, Seitz HK, editors. Alcohol and Cancer. Advances in Experimental Medicine and Biology. Volume 1032. Springer International Publishing; 2018. pp. 127–43. 10.1007/978-3-319-98788-0_10.
2. Krupenko SA, Horita DA. The Role of Single-Nucleotide Polymorphisms in the Function of Candidate Tumor Suppressor ALDH1L1. Front Genet. 2019;10:1013. 10.3389/fgene.2019.01013.
3. Krupenko SA, Krupenko NI. Loss of ALDH1L1 folate enzyme confers a selective metabolic advantage for tumor progression. Chem Biol Interact. 2019;302:149–55. 10.1016/j.cbi.2019.02.013.
4. Krupenko NI, Holmes RS, Tsybovsky Y, Krupenko SA. Aldehyde dehydrogenase homologous folate enzymes: Evolutionary switch between cytoplasmic and mitochondrial localization. Chem Biol Interact. 2015;234:12–7. 10.1016/j.cbi.2014.12.022.
5. Dmitriev AA, Kashuba VI, Haraldson K, et al. Genetic and epigenetic analysis of non-small cell lung cancer with NotI-microarrays. Epigenetics. 2012;7(5):502–13. 10.4161/epi.19801.
6. Chen XQ, He JR, Wang HY. Decreased expression of ALDH1L1 is associated with a poor prognosis in hepatocellular carcinoma. Med Oncol. 2012;29(3):1843–9. 10.1007/s12032-011-0075-x.
7. Rodriguez FJ, Giannini C, Asmann YW, et al. Gene Expression Profiling of NF-1-Associated and Sporadic Pilocytic Astrocytoma Identifies Aldehyde Dehydrogenase 1 Family Member L1 (ALDH1L1) as an Underexpressed Candidate Biomarker in Aggressive Subtypes. J Neuropathol Exp Neurol. 2008;67(12):1194–204. 10.1097/NEN.0b013e31818fbe1e.
8. Ryu HS, Abdellatef AA, Kim DS, Nikas IP, Krupenko SA. Clinical implications of 10-formyltetrahydrofolate dehydrogenase expression in hormone receptor-positive breast cancer. Front Oncol. 2026;16:1838093. 10.3389/fonc.2026.1838093.
9. Zhu G, Liao X, Han C, et al. ALDH1L1 variant rs2276724 and mRNA expression predict post-operative clinical outcomes and are associated with TP53 expression in HBV-related hepatocellular carcinoma. Oncol Rep. 2017;38(3):1451–63. 10.3892/or.2017.5822.
10. Oleinik NV, Krupenko NI, Krupenko SA. Epigenetic Silencing of ALDH1L1, a Metabolic Regulator of Cellular Proliferation, in Cancers. Genes Cancer. 2011;2(2):130–9. 10.1177/1947601911405841.
11. Hoeferlin LA, Fekry B, Ogretmen B, Krupenko SA, Krupenko NI. Folate Stress Induces Apoptosis via p53-dependent de Novo Ceramide Synthesis and Up-regulation of Ceramide Synthase 6. J Biol Chem. 2013;288(18):12880–90. 10.1074/jbc.M113.461798.
12. Greene CL, Jaatinen KJ, Wang H, Koyano TK, Bilbao MS, Woo YJ. Transcriptional Profiling of Normal, Stenotic, and Regurgitant Human Aortic Valves. Genes. 2020;11(7):789. 10.3390/genes11070789.
13. Wu L, Lu X, Guo J, Zhang T, Wang F, Bao Y. Association between ALDH1L1 gene polymorphism and neural tube defects in the Chinese Han population. Neurol Sci. 2016;37(7):1049–54. 10.1007/s10072-016-2527-8.

14. Górska AM, Santos-García I, Eiriz I, Brüning T, Nyman T, Pahnke J. Evaluation of cerebrospinal fluid (CSF) and interstitial fluid (ISF) mouse proteomes for the validation and description of Alzheimer's disease biomarkers. J Neurosci Methods. 2024;411:110239. 10.1016/j.jneumeth.2024.110239.
15. Naz N, Naqvi SF, Hohn N, et al. Cerebral Folate Metabolism in Post-Mortem Alzheimer's Disease Tissues: A Small Cohort Study. Int J Mol Sci. 2022;24(1):660. 10.3390/ijms24010660.
16. Murphy SK, Yang H, Moylan CA, et al. Relationship Between Methylome and Transcriptome in Patients With Nonalcoholic Fatty Liver Disease. Gastroenterology. 2013;145(5):1076–87. 10.1053/j.gastro.2013.07.047.
17. Krupenko SA, Cole SA, Hou R, et al. Genetic variants in ALDH1L1 and GLDC influence the serine-to-glycine ratio in Hispanic children. Am J Clin Nutr. 2022;116(2):500–10. 10.1093/ajcn/nqac091.
18. Stevens VL, McCullough ML, Pavluck AL, et al. Association of Polymorphisms in One-Carbon Metabolism Genes and Postmenopausal Breast Cancer Incidence. Cancer Epidemiol Biomarkers Prev. 2007;16(6):1140–7. 10.1158/1055-9965.EPI-06-1037.
19. Xie W, Wood AR, Lyssenko V, et al. Genetic Variants Associated With Glycine Metabolism and Their Role in Insulin Sensitivity and Type 2 Diabetes. Diabetes. 2013;62(6):2141–50. 10.2337/db12-0876.
20. Williams SR, Yang Q, Chen F et al. Genome-Wide Meta-Analysis of Homocysteine and Methionine Metabolism Identifies Five One Carbon Metabolism Loci and a Novel Association of ALDH1L1 with Ischemic Stroke. Scheet P, ed. *PLoS Genet*. 2014;10(3Jung):e1004214. 10.1371/journal.pgen.1004214
21. Wu IW, Tsai TH, Lo CJ, et al. Discovering a trans-omics biomarker signature that predisposes high risk diabetic patients to diabetic kidney disease. Npj Digit Med. 2022;5(1):166. 10.1038/s41746-022-00713-7.
22. Tabakoglu NT, Ay A, Alkanli N, Celik M. Genotype Combinations and Genetic Risk Score Analyses of MTHFR, MTRR, and MTR Polymorphisms in Hypothyroidism Susceptibility: A Case–Control Study. Curr Issues Mol Biol. 2025;47(10):794. 10.3390/cimb47100794.
23. Bhatia G, Gusev A, Loh PR, et al. Haplotypes of common SNPs can explain missing heritability of complex diseases. Genetics Preprint posted online July. 2015;12. 10.1101/022418.
24. Khankhanian P, Gourraud PA, Lizee A, Goodin DS. Haplotype-based approach to known MS-associated regions increases the amount of explained risk. J Med Genet. 2015;52(9):587–94. 10.1136/jmedgenet-2015-103071.
25. Shim H, Chun H, Engelman CD, Payseur BA. Genome-wide association studies using single-nucleotide polymorphisms versus haplotypes: an empirical comparison with data from the North American Rheumatoid Arthritis Consortium. *BMC Proc*. 2009;3(S7):S35. 10.1186/1753-6561-3-S7-S35
26. Nieman DC, Shanely RA, Luo B, Meaney MP, Dew DA, Pappan KL. Metabolomics approach to assessing plasma 13- and 9-hydroxy-octadecadienoic acid and linoleic acid metabolite responses to 75-km cycling. Am J Physiol-Regul Integr Comp Physiol. 2014;307(1):R68–74. 10.1152/ajpregu.00092.2014.
27. Wang H, Lu J, Kulkarni S, et al. Metabolic and oncogenic adaptations to pyruvate dehydrogenase inactivation in fibroblasts. J Biol Chem. 2019;294(14):5466–86. 10.1074/jbc.RA118.005200.
28. Butte NF, Liu Y, Zakeri IF, et al. Global metabolomic profiling targeting childhood obesity in the Hispanic population. Am J Clin Nutr. 2015;102(2):256–67. 10.3945/ajcn.115.111872.
29. The Multiple Tissue Human Expression Resource (MuTHER) Consortium, Shin SY, Fauman EB, et al. An atlas of genetic influences on human blood metabolites. Nat Genet. 2014;46(6):543–50. 10.1038/ng.2982.
30. Lane CR, Beales PA, Hughes KJD, editors. Fungal Plant Pathogens: Applied Techniques. 2nd ed. CABI; 2023. 10.1079/9781800620575.0000.
31. Dang X, Chen Y, Islam SS, Heath DD, Dixon B. QuantStudio™ 12K flex OpenArray™ chip as a tool for high-throughput immunogene expression analysis in chinook salmon Oncorhynchus tshawytscha. Comp Immunol Rep. 2026;10:200276. 10.1016/j.cirep.2026.200276.
32. Almasy L, Blangero J. Multipoint Quantitative-Trait Linkage Analysis in General Pedigrees. Am J Hum Genet. 1998;62(5):1198–211. 10.1086/301844.

33. Jung J, Sun B, Kwon D, Koller DL, Foroud TM. Allelic-based gene-gene interaction associated with quantitative traits. Genet Epidemiol. 2009;33(4):332–43. 10.1002/gepi.20385.
34. Han CW, Lee HN, Jeong MS, Kim HY, Jang SB. Structural identification and comprehension of human ALDH1L1-Gossypol complex. Biochem Biophys Res Commun. 2024;726:150306. 10.1016/j.bbrc.2024.150306.
35. Krupenko NI, Sharma J, Pediaditakis P, et al. Cytosolic 10-formyltetrahydrofolate dehydrogenase regulates glycine metabolism in mouse liver. Sci Rep. 2019;9(1):14937. 10.1038/s41598-019-51397-1.
36. Fu G, Molina S, Krupenko SA, Sumner S, Rushing BR. Untargeted Metabolomics Reveals Dysregulation of Glycine- and Serine-Coupled Metabolic Pathways in an ALDH1L1-Dependent Manner In Vivo. Metabolites. 2024;14(12):696. 10.3390/metabo14120696.
37. Borecki IB, Province MA. Genetic and Genomic Discovery Using Family Studies. Circulation. 2008;118(10):1057–63. 10.1161/CIRCULATIONAHA.107.714592.
38. Almasy L. Family studies in the age of big data. Proc Natl Acad Sci. 2022;119(15):e2200472119. 10.1073/pnas.2200472119.
39. Guan J, Tan T, Nehzati SM, et al. Family-based genome-wide association study designs for increased power and robustness. Nat Genet. 2025;57(4):1044–52. 10.1038/s41588-025-02118-0.
40. Davies NM, Hemani G, Neiderhiser JM, et al. The importance of family-based sampling for biobanks. Nature. 2024;634(8035):795–803. 10.1038/s41586-024-07721-5.
41. Laird NM, Lange C. The Role of Family-Based Designs in Genome-Wide Association Studies. Stat Sci. 2009;24(4). 10.1214/08-STS280.
42. Dharuri H, Demirkan A, Van Klinken JB, et al. Genetics of the human metabolome, what is next? Biochim Biophys Acta BBA - Mol Basis Dis. 2014;1842(10):1923–31. 10.1016/j.bbadis.2014.05.030.
43. Hagenbeek FA, Pool R, Van Dongen J, et al. Heritability estimates for 361 blood metabolites across 40 genome-wide association studies. Nat Commun. 2020;11(1):39. 10.1038/s41467-019-13770-6.
44. Chen Y, Lu T, Pettersson-Kymmer U, et al. Genomic atlas of the plasma metabolome prioritizes metabolites implicated in human diseases. Nat Genet. 2023;55(1):44–53. 10.1038/s41588-022-01270-1.
45. Underwood MI, Ozel AB, Deepak T, et al. Genome-wide association and linkage analysis of histidine-rich glycoprotein identifies common variants associated with plasma histidine-rich glycoprotein concentrations. Res Pract Thromb Haemost. 2025;9(5):102955. 10.1016/j.rpth.2025.102955.
46. Jia Q, Han Y, Huang P, et al. Genetic Determinants of Circulating Glycine Levels and Risk of Coronary Artery Disease. J Am Heart Assoc. 2019;8(10):e011922. 10.1161/JAHA.119.011922.
47. Whitmore KV, Gaspar HB. Adenosine Deaminase Deficiency – More Than Just an Immunodeficiency. Front Immunol. 2016;7. 10.3389/fimmu.2016.00314.
48. Daddona PE, Kelley WN. Human adenosine deaminase. Purification and subunit structure. J Biol Chem. 1977;252(1):110–5. 10.1016/S0021-9258(17)32805-3.
49. Brosnan ME, Brosnan JT. Histidine Metabolism and Function. J Nutr. 2020;150:S2570–5. 10.1093/jn/nxaa079.
50. Huang H, Li Y, Liang J, Finkelman FD. Molecular Regulation of Histamine Synthesis. Front Immunol. 2018;9:1392. 10.3389/fimmu.2018.01392.
51. Rushing BR, Fogle HM, Sharma J, et al. Exploratory Metabolomics Underscores the Folate Enzyme ALDH1L1 as a Regulator of Glycine and Methylation Reactions. Molecules. 2022;27(23):8394. 10.3390/molecules27238394.
52. Moro J, Tomé D, Schmidely P, Demersay TC, Azzout-Marniche D. Histidine: A Systematic Review on Metabolism and Physiological Effects in Human and Different Animal Species. Nutrients. 2020;12(5):1414. 10.3390/nu12051414.
53. Louck LE, Klatt KC, Wallace TC, Ma J, Chung M. A Genome-Wide Association Study of Circulating Serum Choline, Betaine, Dimethylglycine, and Their Ratios. Nutrients. 2025;17(16):2630. 10.3390/nu17162630.
54. Magnusson M, Wang TJ, Clish C, et al. Dimethylglycine Deficiency and the Development of Diabetes. Diabetes. 2015;64(8):3010–6. 10.2337/db14-1863.

55. Sasaki M, Yamamoto K, Ueda T, et al. One-carbon metabolizing enzyme ALDH1L1 influences mitochondrial metabolism through 5-aminoimidazole-4-carboxamide ribonucleotide accumulation and serine depletion, contributing to tumor suppression. Sci Rep. 2023;13(1):13486. 10.1038/s41598-023-38142-5.
56. Sharma J, Krupenko NI, Krupenko SA. Impact of *Aldh1l1* Knockout On Metabolic Phenotype in Mouse Liver. FASEB J. 2019;33(1). 10.1096/fasebj.2019.33.1_supplement.lb249.
57. Li F, Zhang J, Kang J, Tong X, Zhang P, Zhu Y. Inosine: biofunctions and the roles in human diseases. Front Endocrinol. 2025;16:1634814. 10.3389/fendo.2025.1634814.
58. Srinivasan S, Torres AG, Ribas De Pouplana L. Inosine in Biology and Disease. Genes. 2021;12(4):600. 10.3390/genes12040600.
59. Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. Nat Rev Genet. 2005;6(2):95–108. 10.1038/nrg1521.
60. Balding DJ. A tutorial on statistical methods for population association studies. Nat Rev Genet. 2006;7(10):781–91. 10.1038/nrg1916.

Tables

Table 1: Descriptive characteristics of study participants (n = 409)

Variable	Mean (SD)	H2r(SE)	P-value
Age	23.31 (15.46)		
BMI	24.20 (7.96)	0.61 (0.09)	2.7E-11
Sarcosine	-0.10 (0.47)	1.00 (0.5)	4.5E-36
Cysteine	-0.01 (0.41)	0.64 (0.11)	2.2 E-09
N. formylmethionine	0.01 (0.21)	0.56 (0.10)	5.3 E-09
AMP	-0.03 (0.61)	0.47 (0.09)	2.7 E-08
Glycine	0.00 (0.25)	0.55 (0.11)	3.4 E-08
Hypoxanthine	-0.02 (0.58)	0.56 (0.11)	4.9 E-08
Pantothenate	0.01 (0.36)	0.48 (0.10)	1.0 E-07
Hippurate	-0.03 (0.77)	0.44 (0.10)	4.0 E-07
Creatine	-0.08 (0.46)	0.53 (0.11)	6.0 E-07
Xanthosine	-0.02 (0.38)	0.40 (0.10)	3.7 E-06
Serine	-0.03 (0.22)	0.43 (0.11)	3.9 E-06
Dimethylglycine	0.00 (0.33)	0.41 (0.10)	3.2 E-05
Urate	0.00 (0.25)	0.42 (0.11)	3.3 E-05
N. acetylglycine	0.06 (0.48)	0.36 (0.10)	4.2 E-05
Histidine	0.01 (0.19)	0.34 (0.10)	1.2 E-04
Citrulline	-0.04 (0.29)	0.36 (0.11)	1.4 E-04
Serotonin	-0.57 (1.06)	0.30 (0.09)	2.1 E-04
Cinnamoylglycine	-0.33 (1.41)	0.26 (0.10)	2.0 E-03
Inosine	-0.16 (0.59)	0.28 (0.10)	2.2 E-03
Creatinine	-0.01 (0.21)	0.26 (0.10)	2.6 E-03
Xanthine	-0.03 (0.37)	0.24 (0.10)	2.8 E-03
Uridine	-0.01 (0.23)	0.27 (0.11)	3.6 E-03
Formiminoglutamate	-0.03 (0.57)	0.26 (0.11)	5.5 E-03
Ornithine	-0.50 (1.19)	0.21 (0.10)	1.2 E-02
Urea	0.02 (0.34)	0.22 (0.10)	1.4 E-02
Myoinositol	-0.00 (0.28)	0.18 (0.11)	3.4 E-02
Adenosine	-0.25 (0.57)	0.13 (0.09)	7.4 E-02
Inosine/Hypoxanthine	-1.42 (31.77)	0.001 (0.09)	0.50
Adenosine/Inosine	-1.44 (51.88)	0.01 (0.08)	0.43
Serine/Glycine	1.96 (28.32)	0.001 (0.09)	0.50

Table 2: Six exonic variants of *ALDH1L1* utilized in this study

a. Genomic positions and frequencies of the six *ALDH1L1* SNPs

SNP	Allele (Major/Minor)	Frequency (%)	HWE* (p value)
rs1127717	T/C	19.1	0.3
rs2276724	T/C	17.5	0.2
rs2886059	C/A	22	0.8
rs3796191	A/G	4.3	0.5
rs4646750	T/C	10.2	0.6
rs9282691	T/G	4	1.0

*SNP – Single Nucleotide Polymorphisms; HWE – Hardy-Weinberg Equilibrium

b. Association of *ALDH1L1* SNPs with glycine and purine metabolites

SNP	Metabolite	Estimate (SE)	P-value
rs1127717	n_formylmethionine	0.137 (0.051)	7.6E-03
	myoinositol	0.277 (0.068)	6.1E-05
	urea	0.201 (0.083)	1.6E-02
rs2276724	histidine	-0.051 (0.021)	1.5E-02
	Hippurate	-0.236 (0.086)	6.1E-03
	Cinnamoylglycine	-0.563 (0.155)	3.3E-04
rs2886059	histidine	-0.049 (0.020)	1.3E-02
	Cinnamoylglycine	-0.335 (0.148)	2.5E-02
rs3796191	hypoxanthine	-0.217 (0.106)	4.1E-02
	glycine	0.110 (0.046)	1.6E-02
	histidine	0.139 (0.033)	3.2E-05
	inosine	-0.225 (0.106)	3.5E-02
	pantothenate	0.183 (0.065)	5.2E-03
	adenosine_inosine ratio	-22.741 (9.450)	1.7E-02
rs4646750	inosine	0.181 (0.075)	1.7E-02
	serine	0.086 (0.028)	2.1E-03
	sarcosine	-0.269 (0.060)	9.7E-06
rs9282691	methionine	0.074 (0.031)	1.7E-02
	pantothenate	-0.150 (0.065)	2.2E-02

Table 3: Genotypic combinations of the six *ALDH1L1* exonic SNPs (Combinations > 2% frequency shown here)

Genotypic Combination	Frequency	Percentage (%)
rs1127717*TT_rs2276724*TT_rs2886059*CC_rs3796191*AA_rs4646750*TT_rs9282691*TT	129	31.62
rs1127717*CT_rs2276724*TT_rs2886059*CC_rs3796191*AA_rs4646750*TT_rs9282691*TT	57	13.97
rs1127717*TT_rs2276724*CT_rs2886059*AC_rs3796191*AA_rs4646750*TT_rs9282691*TT	35	8.58
rs1127717*TT_rs2276724*CT_rs2886059*AC_rs3796191*AA_rs4646750*CT_rs9282691*TT	29	7.11
rs1127717*CT_rs2276724*CT_rs2886059*AC_rs3796191*AA_rs4646750*TT_rs9282691*TT	25	6.13
rs1127717*TT_rs2276724*TT_rs2886059*CC_rs3796191*AG_rs4646750*TT_rs9282691*TT	20	4.9
rs1127717*CT_rs2276724*TT_rs2886059*CC_rs3796191*AA_rs4646750*CT_rs9282691*TT	10	2.45
rs1127717*TT_rs2276724*TT_rs2886059*AC_rs3796191*AA_rs4646750*TT_rs9282691*GT	9	2.21
rs1127717*TT_rs2276724*TT_rs2886059*CC_rs3796191*AA_rs4646750*CT_rs9282691*TT	9	2.21

Table 4: Association of genotypic combinations of exonic *ALDH1L1* SNPs with glycine and purine metabolites*

Genotypic Combination	Metabolite	Estimate (SE)	P-value
rs1127717*TT_rs2276724*CT_rs2886059*AC_ rs3796191*AG_rs4646750*TT_rs9282691*TT	Adenosine/Inosine	-191.005 (24.529)	5.9E-14
rs1127717*TT_rs2276724*TT_rs2886059*CC_ rs3796191*AG_rs4646750*TT_rs9282691*TT	Histidine	0.219 (0.041)	2.1E-07
rs1127717*TT_rs2276724*TT_rs2886059*CC_ rs3796191*AA_rs4646750*TT_rs9282691*GT	Adenosine/Inosine	-123.824 (25.497)	1.7E-06
rs1127717*TT_rs2276724*TT_rs2886059*AC_ rs3796191*AA_rs4646750*TT_rs9282691*GT	Dimethylglycine	0.510 (0.109)	3.7E-06
rs1127717*TT_rs2276724*CC_rs2886059*AA_ rs3796191*AA_rs4646750*CT_rs9282691*TT	Adenosine/Inosine	99.980 (23.004)	1.8E-05
rs1127717*TT_rs2276724*TT_rs2886059*AC_ rs3796191*AA_rs4646750*CT_rs9282691*TT	Inosine	1.412 (0.333)	2.8E-05
rs1127717*TT_rs2276724*CT_rs2886059*AC_ rs3796191*AA_rs4646750*TT_rs9282691*TT	Serine	-0.148 (0.038)	1.1E-04

*p value adjusted for the number of tests is 7.1 E-4

Table 5: Association of allelic combinations of *ALDH1L1* SNPs with glycine metabolites

Trait	SNPs	Estimate	StdError	P value
adenosine_inosine_ratio	rs2276724*C_rs3796191*G	-638.44	95.38	7.4E-11
adenosine_inosine_ratio	rs2886059*A_rs3796191*G	-461.89	84.89	9.3E-08
histidine	rs2886059*A_rs3796191*A	-0.32	0.07	9.9E-06
histidine	rs2276724*C_rs3796191*A	-0.31	0.07	1.3E-05
Cinnamoylglycine	rs1127717*T_rs2276724*C	-1.33	0.31	1.5E-05
histidine	rs1127717*T_rs3796191*A	-0.29	0.07	1.9E-05
histidine	rs3796191*A_rs4646750*T	-0.29	0.07	2.7E-05
histidine	rs3796191*A_rs4646750*C	-0.31	0.08	3.5E-05
histidine	rs2276724*T_rs3796191*A	-0.28	0.07	3.8E-05
histidine	rs2886059*C_rs3796191*A	-0.29	0.07	4.3E-05
adenosine_inosine_ratio	rs1127717*T_rs2276724*C_rs3796191*G	-645.20	96.48	7.7E-11
adenosine_inosine_ratio	rs1127717*T_rs2886059*A_rs3796191*G	-469.94	86.19	8.7E-08
adenosine_inosine_ratio	rs2276724*C_rs2886059*A_rs3796191*G	-1242.61	229.17	1.0E-07
histidine	rs1127717*T_rs2886059*A_rs3796191*A	-0.39	0.08	4.5E-07
histidine	rs1127717*T_rs2276724*C_rs3796191*A	-0.38	0.08	7.4E-07
adenosine_inosine_ratio	rs2276724*C_rs3796191*G_rs9282691*T	-669.14	133.63	8.3E-07
adenosine_inosine_ratio	rs2886059*A_rs3796191*G_rs9282691*T	-964.20	193.11	8.9E-07
histidine	rs1127717*C_rs2886059*C_rs3796191*A	-0.35	0.08	1.3E-05
histidine	rs1127717*C_rs2276724*T_rs3796191*A	-0.33	0.08	1.8E-05
Cinnamoylglycine	rs1127717*T_rs2276724*C_rs4646750*T	-1.76	0.41	2.5E-05
histidine	rs1127717*T_rs2276724*T_rs3796191*A	-0.31	0.07	2.5E-05
histidine	rs1127717*T_rs2886059*C_rs3796191*A	-0.31	0.08	3.4E-05
adenosine_inosine_ratio	rs1127717*T_rs2886059*C_rs9282691*G	-194.41	46.59	3.7E-05
pantothenate	rs2276724*T_rs2886059*A_rs3796191*A	-0.68	0.17	6.1E-05
histidine	rs2276724*C_rs2886059*C_rs3796191*A	-0.46	0.11	7.1E-05
adenosine_inosine_ratio	rs2886059*A_rs3796191*G_rs4646750*T_rs9282691*T	-880.50	145.73	3.5E-09
adenosine_inosine_ratio	rs2276724*C_rs3796191*G_rs4646750*T_rs9282691*T	-810.27	140.34	1.6E-08
adenosine_inosine_ratio	rs1127717*T_rs2276724*C_rs2886059*A_rs3796191*G	-1240.95	230.09	1.2E-07
adenosine_inosine_ratio	rs1127717*T_rs2276724*C_rs3796191*G_rs9282691*T	-681.54	134.05	5.7E-07
adenosine_inosine_ratio	rs1127717*T_rs2886059*A_rs3796191*G_rs9282691*T	-946.24	193.03	1.4E-06
histidine	rs1127717*C_rs2276724*T_rs2886059*C_rs3796191*A	-0.33	0.08	2.8E-05
inosine	rs2276724*T_rs2886059*A_rs4646750*C_rs9282691*T	8.11	1.92	3.0E-05
adenosine_inosine_ratio	rs2276724*C_rs2886059*A_rs3796191*G_rs4646750*T	-2579.38	639.83	6.7E-05
histidine	rs1127717*T_rs2276724*T_rs2886059*C_rs3796191*A	-0.30	0.08	7.7E-05
adenosine_inosine_ratio	rs1127717*T_rs2886059*A_rs3796191*G_rs4646750*T	-932.92	143.04	2.2E-10

	_rs9282691*T			
adenosine_inosine_ratio	rs1127717*T_rs2276724*C_rs3796191*G_rs4646750*T _rs9282691*T	-811.07	134.79	4.1E-09
inosine	rs2276724*T_rs2886059*A_rs3796191*A_rs4646750*C _rs9282691*T	8.49	1.86	6.6E-06
adenosine_inosine_ratio	rs1127717*T_rs2886059*C_rs3796191*A_rs4646750*T _rs9282691*G	-464.73	104.52	1.1E-05
adenosine_inosine_ratio	rs2276724*C_rs2886059*A_rs3796191*G_rs4646750*T _rs9282691*T	-1501.87	337.88	1.2E-05
adenosine_inosine_ratio	rs1127717*T_rs2276724*C_rs2886059*A_rs3796191*G _rs4646750*T	-2335.20	567.57	4.8E-05
adenosine_inosine_ratio	rs1127717*T_rs2276724*C_rs2886059*A_rs3796191*G _rs4646750*T_rs9282691*T	-1480.06	336.88	1.5E-05

*rs1127717 - T/C; rs2276724 - T/C; rs2886059 - C/A; rs3796191 - A/G; rs4646750 - T/C; rs9282691 - T/G

**p value adjusted for the number of tests is 7.8 E-5